

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:07 ; Search time 299.73 Seconds

(Without alignments)
260.436 Million cell updates/sec

Title: US-09-331-631A-3

Perfect score: 3532
Sequence: 1 MAINTSNLCSLLFLSLFL.....SSRSTKQOQPLVSIIDFVGF 666

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP:REMBL_15:*
2: SP:archaea:*
3: SP:bacteria:*
4: SP:fungi:*
5: SP:human:*
6: SP:mammal:*
7: SP:mhc:*
8: SP:organelle:*
9: SP:phage:*
10: SP:plant:*
11: SP:rodent:*
12: SP:virus:*
13: SP:vertebrate:*
14: SP:unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3532	100.0	666	10	Q9SP14 macadamia i
2	3412	96.6	666	10	Q9SP15 macadamia i
3	3215	91.0	625	10	Q9SP13 macadamia i
4	1324.5	37.5	593	10	Q9SEW4 juglans reg
5	1191.5	33.7	810	10	Q9ZWI3 cucurbita m
6	1036	29.3	525	10	Q43358 theobroma c
7	895	25.3	582	10	Q03865 zea mays (m
8	875	24.8	637	10	Q03678 hordeum vul
9	851	24.1	613	10	Q9M3X6 pisum sativ
10	833	23.6	544	10	Q22120 glycine max
11	818	23.2	545	10	Q41674 vicia narbo
12	798.5	22.6	489	10	Q9SP11 glycine max
13	770.5	21.6	483	10	Q49927 pisum sativ
14	765	21.4	448	10	Q40873 pisca glauc
15	757	21.4	417	10	Q22121 glycine max
16	754.5	21.4	463	10	Q41677 vicia narbo
17	754	21.3	450	10	Q40844 pisca glauc
18	749.5	21.3	438	10	Q43626 pisum sativ
19	746	21.1	486	10	Q9JL07 arabidopsis

20	706	20.0	518	10	Q9M3X8 lens culina
21	635	18.0	461	10	Q9ZRG9 oryza sativ
22	565	16.0	430	10	Q41115 phaseolus v
23	561	15.9	430	10	Q43633 phaseolus v
24	554	15.7	421	10	Q43632 phaseolus v
25	552.5	15.6	414	10	Q41727 zania furfu
26	541.5	15.3	232	10	Q9SQ49 rullingia ma
27	534.5	15.1	540	10	Q03866 zea mays (m
28	527	14.9	239	10	Q9SQ50 abroma augu
29	527	14.9	423	10	Q43617 phaseolus l
30	523.5	14.8	239	10	Q40913 phaseolus l
31	522	14.8	239	10	Q9SQ42 theobroma b
32	521.5	14.8	236	10	Q9SQ41 theobroma c
33	521	14.8	238	10	Q9SQ32 theobroma s
34	520.5	14.7	236	10	Q9SQ47 theobroma s
35	520	14.7	239	10	Q9SQ44 theobroma v
36	520	14.7	239	10	Q9SQ35 theobroma v
37	520	14.7	239	10	Q9SV79 theobroma c
38	518.5	14.7	236	10	Q9SQ40 theobroma c
39	517.5	14.7	236	10	Q9SQ36 theobroma c
40	517.5	14.7	236	10	Q9SQ34 theobroma g
41	516.5	14.6	236	10	Q9SQ37 theobroma g
42	515.5	14.6	236	10	Q9SQ48 theobroma m
43	515	14.6	239	10	Q9SQ43 theobroma m
44	510.5	14.5	236	10	Q9SQ33 theobroma m
45	508	14.4	235	10	Q9SQ38 theobroma g

ALIGNMENTS

RESULT 1
Q9SP14 PRELIMINARY: PRT: 666 AA.
ID Q9SP14;
AC Q9SP14; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OC NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUIT KERNEL.
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia.";
RL Plant J. 0:0-0(1999).
DR EMBL: AF161884; AMD54245.1; -
DR HSSP: P02853; 2PH.
DR INTERPRO: IPR001113; -
DR PFM: PF00546; Seedscore_7s; 1.
DR PFM: PF00546; Seedscore_7s; 1.
SQ SEQUENCE 666 AA: 78243 MW; 0ECA22P8710F8A7B CRC64;

Query Match 100.0%; Score 3532; DB 10; Length 666;
Best Local Similarity 100.0%; Pred. No. 2e-248;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAINTSNLCSLLFLSLFLSTVSLAESFPDQFYEBCRKQCMOLETSQMRVCVSOCD 60
DB 1 MAINTSNLCSLLFLSLFLSTVSLAESFPDQFYEBCRKQCMOLETSQMRVCVSOCD 60
QY 61 KREEDIDMSKYDNDPQDPCQCORRCQOESGPRQOYCORCKEICEEEYRNOR 120
DB 61 KREEDIDMSKYDNDPQDPCQCORRCQOESGPRQOYCORCKEICEEEYRNOR 120
QY 121 DPQOYEQCOERQORHETFERHMQTCOQRCERYEKKRQKRYEQOQREDEKYEEM 180
DB 121 DPQOYEQCOERQORHETFERHMQTCOQRCERYEKKRQKRYEQOQREDEKYEEM 180

QY	181	KEEDKRPQOREVEDCRRREEOEPPOOVYOCORCEQOOROHGRGDDLINPORGSGRY	240
QY	181	KEEDKRPQOREVEDCRRREEOEPPOOVYOCORCEQOOROHGRGDDLINPORGSGRY	240
Db	181	KEEDKRPQOREVEDCRRREEOEPPOOVYOCORCEQOOROHGRGDDLINPORGSGRY	240
QY	241	EEGEEKOSDNFYEDERSLSTRFTTEEGHISVLENFYGRSKLRALKNYRLVLEANPNA	300
Db	241	EEGEEKOSDNFYEDERSLSTRFTTEEGHISVLENFYGRSKLRALKNYRLVLEANPNA	300
QY	301	FVLPHLDADAILLVGTGRGALKMHRDNRSEYNLECGDVIRIPAGTFYFLINRDNREL	360
Db	301	FVLPHLDADAILLVGTGRGALKMHRDNRSEYNLECGDVIRIPAGTFYFLINRDNREL	360
QY	361	HIKFLQITSTPGQYKEFFPAGGONPEPYLSTFESKELLEALNTQAERLNGVLGOOREGV	420
Db	361	HIKFLQITSTPGQYKEFFPAGGONPEPYLSTFESKELLEALNTQAERLNGVLGOOREGV	420
QY	421	IISASOEQIRELTITDSDSRKWHIRRGESSRGCPYNLFNKRPYLNKYGQAYEVKPEDYR	480
Db	421	IISASOEQIRELTITDSDSRKWHIRRGESSRGCPYNLFNKRPYLNKYGQAYEVKPEDYR	480
QY	481	QLODMDSVFLANTTOSGMGMPFNTSTKVYVVASGEADVEMACPHLSGRHGRGRGKR	540
Db	481	QLODMDSVFLANTTOSGMGMPFNTSTKVYVVASGEADVEMACPHLSGRHGRGRGKR	540
QY	541	HEEEDVHYEQVKARLSKREAIIVPVGHPVYVSSGNENLLEAFGINAONHNEFLAGR	600
Db	541	HEEEDVHYEQVKARLSKREAIIVPVGHPVYVSSGNENLLEAFGINAONHNEFLAGR	600
QY	601	ERNVLOQIEPQAMELAPAPKVEELFNODESIFPPGRPHOQOSRSTKQOQPLVSI	660
Db	601	ERNVLOQIEPQAMELAPAPKVEELFNODESIFPPGRPHOQOSRSTKQOQPLVSI	660
QY	661	LDPVGF 666	
Db	661	LDPVGF 666	
RESULT	2		
Q9SP15	PRELIMINARY;	PRU;	666 AA.
AC	09SP15		
DT	01-MAY-2000 (Tremblere1, 13, Created)		
DT	01-MAY-2000 (Tremblere1, 13, Last sequence update)		
DT	01-OCT-2000 (Tremblere1, 15, Last annotation update)		
DE	VICILIN PRECURSOR.		
GN	AMP2.		
OS	Macadamia integrifolia (Macadamia nut).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Manoliphyta; eudicotyledons; Proteaceae; Macadamia.		
OX	NCBI_TaxId=60698;		
RN	11		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=NUIT KERNEL;		
RA	Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;		
RT	"A family of antimicrobial peptides is produced by processing of a 7S		
RT	globulin protein in Macadamia integrifolia kernels."		
RL	Plant J. 0:0-0(1999).		
DR	EMBL: AF161883; AAD54244.1; -		
DR	HSSP: P02853; 2PHL.		
DR	INTERPRO: IPR001113; -		
DR	PFAM: PF00546; Seedstore_7s; 1.		
DO	SEQUENCE 666 AA; 78217 MW; C752B884B2DF0224 CRC64;		

```

Query Match      96.6%  Score 3412:  DB 10:  Length 666:
Best Local Similarity 96.1%:  Pred No 1e-229:
Matches 640:  Conservative 12:  Mismatches 14:  Indels 0:  Gaps 0

OY 1 MAINTSNLCSTLFLSLSTFLSTTVSLAESEFDRDEYECKRCOMOLETSGOMKRCVSGCD 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
ob 1 MAINSNLCSTLFLSLSTFLSTTVSLAESEFDRDEYECKRCOMOLETSGOMKRCVSGCD 60

```

QY	61	KRFEEIDMSYDNDQDPOTCCQCCQRCRROEGSPROOQYCCQRCKEICEEPEEYROR	120
Db	61	KRFEEIDMSYDNDQDPOTCCQCCQRCRROEGSPROOQYCCQRCKEICEEPEEYROR	120
QY	121	DPQOQYEOCQERQHETEPRHMTQOQRCERRYEKERRKQOQRYEEOQREDEEYER	180
Db	121	DPQOQYEOCQERQHETEPRHMTQOQRCERRYEKERRKQOQRYEEOQREDEEYER	180
QY	181	KEEENKRDPOQREYEDCRRREDOEPQOQOQOLNCEQOQOJHGRGGDMNPOQGGGRY	240
Db	181	KEEENKRDPOQREYEDCRRREDOEPQOQOQOLNCEQOQOJHGRGGDMNPOQGGGRY	240
QY	241	EEGEEKOSDNPYFEDERSLSTRFTBEGHLSVLENFYGRSKLRLALNRYLVLEANPNA	300
Db	241	EEGEEKOSDNPYFEDERSLSTRFTBEGHLSVLENFYGRSKLRLALNRYLVLEANPNA	300
QY	301	FVLFTHLDADAILLVYTGGRGALKMTHNDNRESYNLECGDVIRIPAGTFFYLINDNNE	360
Db	301	FVLFTHLDADAILLVYTGGRGALKMTHNDNRESYNLECGDVIRIPAGTFFYLINDNNE	360
QY	361	HIKFLQTIISPGQYKEFFPAGGONPEPYLSTFSKELTEALNTQARLKGVLQOQREGV	420
Db	361	HIKFLQTIISPGQYKEFFPAGGONPEPYLSTFSKELTEALNTQKLGVEEQOQREGV	420
QY	421	IISASQEOIRELTRDQESRRMHITRGGESSRGPNLEFNRPYLSNKYGAAYEVKPEDYR	480
Db	421	IIRASQEOIRELTRDQESRRMHITRGGESSRGPNLEFNRPYLSNKYGAAYEVKPEDYR	480
QY	481	QLOQMDVSVFLANTQOSMMGPFNTSTKYVVVAVASEADYEMACPLSLRHGGRGGR	540
Db	481	QLOQMDVSVFLANTQOSMMGPFNTSTKYVVVAVASEADYEMACPLSLRHGGRGGR	540
QY	541	HEEEDVHYEDVKARLSKREAIIVPVGHVYSSGNNENLLFAFGINAQNNHNFAGR	600
Db	541	HEEEDVHYEDVKARLSKREAIIVPVGHVYSSGNNENLLFAFGINAQNNHNFAGR	600
QY	601	ERNVLOQIEPQAMBLAPAPKKEVEELFNSODESIFPPGPRQOQOOSKSTKQOQPLVSI	660
Db	601	ERNVLOQIEPQAMBLAPAPKKEVEELFNSODESIFPPGPRQOQOOSKSTKQOQPLVSI	660
QY	661	LDPFVG 666	
Db	661	LDPFVG 666	
RESULT	3		
Q9SPJ3		PRELIMINARY:	PRT: 625 AA.
AC	Q9SPJ3		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-OCT-2000 (Tremblrel. 15, Last annotation update)		
DE	VICILIN PRECURSOR (FRAGMENT).		
OS	Macadamia integrifolia (Macadamia nut).		
OC	Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta;		
OC	Magnoliophyta: eudicotyledons: Proteaceae: Macadamia.		
OX	NCBI_TaxID=60698;		
FN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=NUIT KERNEL:		
RT	Marcus J.P., Goulier K.C., Green J.L., Manners J.M.:		
RT	"A family of antimicrobial peptides is produced by processing of a 7S		
RT	globulin protein in Macadamia integrifolia."		
RL	Plant J. 0:0-0(1999).		
DR	EMBL: AF161885; AAD54246.1; -.		
DR	HSP: P02853; 2PHL.		
DR	INTERPRO: IPR001113; -.		
DR	PRAM: PF00546; Seedstore_7s; 1.		
FT	NON_TER		
QO	SEQUENCE 625 AA; 73586 MW; 415808A89D370296 CRC64;		

DR EMBL; AF161885; AAD54246.1; -.
DR HSSP; P02853; 2PHL.
DR INTERPRO; IPR001113; -.
DR PFM; PF00546; Seedstore_7s; 1.
FT NON_TER 1 1
SQ SEQUENCE 625 AA; 73386 MW; 415808A89D370296 CRC64

Query Match	91.0%:	Score 3215:	DB 10:	Length 625:
Best Local Similarity	96.6%:	Pred. No. 1.9e-225:		
Matches 604:	Conservative 7:	Mismatches 14:	Indels 0:	Gaps 0:
OY	42	QCMQLETSQOMRRCYSOCCKRFEEDIDMSKYNOQDDPQTDCCQCCORRCQOEESGPRQOY	101	
DB	1	QCMQLETSQOMRRCYSOCCKRFEEDIDMSKYNOQDDPQTDCCQCCORRCQOEESGPRQOY	60	
OY	102	CORCKEICEEEBEEYNRODPQOYECOCORHETEPHMHOTQOOCERRYEKEKRMO	161	
DB	61	CORCKEICEEEBEEYNRODPQOYECOCORHETEPHMHOTQOOCERRYEKEKRMO	120	
OY	162	OKRYEEOQRDEBEKEEENKEDNKRDPOQREYECRCRCEQOEPRQOYQOCRCRCEQR	221	
DB	121	OKRYEEOQRDEBEKEEENKEDNKRDPOQREYECRCRCEQOEPRQOYQOCRCRCEQR	180	
OY	222	QHRCGGDILINPQSGSGRGEEGEESQSDPYPYFDERSILSTRPTEEGHISYLENFGNSK	281	
DB	181	QHRCGGDILINPQSGSGRGEEGEESQSDPYPYFDERSILSTRPTEEGHISYLENFGNSK	240	
OY	282	ILRALKNYRLVYLEANPNAFVLPTHLDADALLVTGGRGALKMIRHNRRESYNLECCDVI	341	
DB	241	ILRALKNYRLVYLEANPNAFVLPTHLDADALLVTGGRGALKMIRHNRRESYNLECCDVI	300	
OY	342	RIPACTTYLLINRDNNERLHAKFLQTTSTPGQYKEFPAGQONPEPLSTFSKILLFAA	401	
DB	301	RIPACTTYLLINRDNNERLHAKFLQTTSTPGQYKEFPAGQONPEPLSTFSKILLFAA	360	
OY	402	LNTQAEERLRGVGGQRECVTISASOEQIRELTRDSESRMHIRRGSSRRPYLWPKNR	461	
DB	361	LNTQAEERLRGVGGQRECVTISASOEQIRELTRDSESRMHIRRGSSRRPYLWPKNR	420	
OY	462	PLYSKNYGOAAYEVKKEPDYROLQDMQVSYFIANTTQGSMMGPEFFNTRSTKVVVVASGEADV	521	
DB	421	PLYSKNYGOAAYEVKKEPDYROLQDMQVSYFIANTTQGSMMGPEFFNTRSTKVVVVASGEADV	480	
OY	522	EMACHLSGRHGGRGRGGRHHEEDVHYEYKARLSKREDAIVYPVGHVYVSSGNENLL	581	
DB	481	EMACHLSGRHGGRGRGGRHHEEDVHYEYKARLSKREDAIVYPVGHVYVSSGNENLL	540	
OY	582	LEAFETINQNNHNFNLACRERNVLOQIIPQAMETLFAFAPRKEVEELTFNSODESTFFPQPR	641	
DB	541	LEAFETINQNNHNFNLACRERNVLOQIIPQAMETLFAFAPRKEVEELTFNSODESTFFPQPR	600	
OY	642	QHQQSSRSRTKQQQPLVLSILDVFCF 666		
DB	601	QHQQSSRSRTKQQQPLVLSILDVFCF 625		
RESULT	4			
Q9SEW4	Q9SEW4	PRELIMINARY:	PRT:	593 AA.
AC	O9SEW4			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	VICILIN-LIKE PROTEIN PRECURSOR (PRAGMENT).			
OS	Juglans regia (English walnut).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I;			
OC	Fagales; Juglandaceae; Juglans.			
OX	NCBI_TaxID=51240;			
RN	111			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV. SUNLAND; TISSUE-SOMATIC EMBRYO LINE;			
RA	Tender S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;			
RT	"Identification and cloning of a cDNA encoding a vicilin-like protein,			
RT	Jug r 2, from English walnut kernel (Juglans regia): a major food			
RT	allergen.";			
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF060605; AAF18269.1; --			
DR	HSSP; P02853; 2PHL.			
DR	INTERPRO; IPR001113; --			

[illegible]

RT of a single precursor by vacuolar processing enzyme."
RL J. Biol. Chem. 274:2563-2570(1999).
DR EMBL: AB019195; BAA34056.1; -
DR HSSP: P02853; 2PHL.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore_7s; 1.
DR PRODOM: PD081059; -; 1.
SQ SEQUENCE 810 AA; 97314 MW; A829A3F7542266AB CRC64;

Query Match 33.7% Score 1191.5; DB 10; Length 810;
Best Local Similarity 34.3%; Pred No. 2.7e-78;
Matches 284; Conservative 144; Mismatches 218; Indels 183; Gaps 23;

```
QY 1 MAINTSNLCLFLFLSTVSLAE-----SEF- 31
DB 1 MALSKYLRCLLAFLLACLSTVGLDGKESLSGAGVDHDCVNRCELKGNVDEFA 60
QY 32 -----DROYECKRQCMQLETS-GQMRVCYSCDKREE----- 65
DB 61 ACKKGGVNRGSPRAEYECRLRCOVAERGVQQRKCEVCEERLEREREGEDVDEV 120
QY 66 ---DIDMSKY-----DNDDPOTDOQCCORRCROE 93
DB 121 ERDDPMEBERORRRRREBERRRRRERGRGRGRGRGRGRGRGRGRGRGRGRGR 178
QY 94 SGFRQOQYQORCKELCEEESEYVNRQDPOQOYEQOCORHETEPRHMQTCQRCERR 153
DB 179 RRRREQERERERGERDEDEDEN-QRDPDMRREQ-----ERREQRRRREQERREQ 232
QY 154 Y-----EKE-----KKQCKRYEQOREDEBEKEEKKEDDKRDP-----QOREY 195
DB 233 RRGGRDEDEENQRPDMRREQERERERERERERERERERERERERERERERER 292
QY 196 DCRRCRQOE-----PROOYOCORCRPOORHGRG-- 227
DB 293 ERRRRERERERERERGRGVDENQRPDMRREQERERERERERERERERERERER 352
QY 228 -----DLINPORGSGRYEE-----GEEKSDNPYFEDRSLSTRTEEGHISVL 273
DB 353 BQSRREDERRRRRQHGSRVQVAVIRTEQGSNNPFFQORQSRQVRSDEGHMRYL 412
QY 274 ENFYGSKLLRALKRVLLENPAFVLPHTLDADALLVTGGGALKMHRDKR--- 330
DB 413 ERSESELLKGIKNQRLALLEARPHFTFVPHLDACVLLVVRGRTITTVQEKRETR 472
QY 331 -SSYNLECGDVIRIPAGTFEYLLINRDNNERLHIAKFLQTIISPGQYKEFFPAGGONPEPY 389
DB 473 KESYNVESGDVMTIPAGTTLTYLANOE-NEDLOIVKLVQVYVNNPGEKDYLSAGES-QAY 530
QY 390 LSTFSKEILLALNTOAERLQVLRGQORE--GVIISASQEOIRELTRDSESRMHIRRG 447
DB 531 YSFVSDVLEALNIPROKLERIFKORRERGGKIVRASQEOIRLALSO-----RATSVRKG 585
QY 448 GSSRRPYLNFKRLPKSKYGOAYEVKPEDYROLQDMVSVFIANIITGOSMMGPFNFR 507
DB 586 SSGVRLPILLESQTPYNNQYQMEACDPDEPQLKRTVAISVDIKGGMVNFENR 645
QY 508 STRVVVVASGEADVACPHLS-----GRHGRGRGKRHEEE-----DVHYEQVARS 557
DB 646 AHWVYFVSGASFEMACHIOSOMQORGRREERHMRREERERERERERERERER 705
QY 558 KREAIIVPVGHVAVVSSGNENLLFAGFINAONNHNENFLARERNVLOQIEPOAMELAF 617
DB 706 EGGVLIYPAGHPLAIMASPNENLRLVGFGINAENNRHNLAGE-NIMMELBREAKELAF 764
QY 618 AAPKREVEELFNSODESITFFPPRHOQOOSRSTKQOQPLVSLIDFVG 666
DB 765 NVEGQADEIFRSORESEFTEGPEGRRRST-----ERSPLLSILKLGY 809
```

RESULT 6
Q43358

ID 043358 PRELIMINARY; PRT; 525 AA.
AC Q43358;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN CSV.
OS Theobroma cacao (Cacao).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Malvales; Malvaceae; Theobroma.
OX NCBI_TaxID=3641;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAVES;
RX MEDLINE=92288309; PubMed=1600151;
RA McHenry L., Filtz P.J.;
RT "Comparison of the structure and nucleotide sequences of vicilin genes
of cocoa and cotton raise questions about vicilin evolution.";
RL Plant Mol. Biol. 18:1173-1176(1992).
DR EMBL: X62625; CAA44493.1; -
DR HSSP: P02853; 2PHL.
DR MENDEL: 30919; Thecc; 1188; 30919.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore_7s; 1.
DR PRODOM: PD081059; -; 1.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 1 24 POTENTIAL.
SQ SEQUENCE 525 AA; 60798 MW; 19114CD5C248905D CRC64;

Query Match 29.3% Score 1036; DB 10; Length 525;
Best Local Similarity 41.4%; Pred No. 2.9e-67;
Matches 213; Conservative 91; Mismatches 156; Indels 54; Gaps 11;

```
QY 109 ICEEEENR---QRDPQOYEQOCORHETEPRHMQTCQRCRRYERKRQOKRY 165
DB 22 LGSVAGYGRKQYENDPRQOYEQOCRCSEATEERREQCQRCR-----EY 70
QY 166 EQQREDEKYEERKMEENKRDPOOREYEDCRRCEQOE--PROOYOCORCRQOROH 223
DB 71 KEOQROQEBEL-----QROYQOCQRCQEOQOQOREQOOCORCMEQYKQO 116
QY 224 GRGGLINPORGSGRYEGEGEEKSDNPYFDE-RSLSTRFTEEGHISVLENYGRSL 282
DB 117 ER-GEHENYHNKKNRSEEEQOQRNNPYFFPKRRSFQTRFRDEBGNFKILQFAENSP 175
QY 283 LRLKNYRLVLEANPAFVLPHTLDADA1LVYGGGALKMHRDNRESYNLECGDVIR 342
DB 176 LKGINYRLAMEANPNFTILPHICDAEAIYVTNGKGTITFEVHNKESYVQRGTVS 235
QY 343 IPAGTFYLLINDNNRHLIAKFLQTIISPGQYKEFFPAGGONPEPYLSTFSKEILLAL 402
DB 236 VPGSTVYVSDNOEKLITAVLALFVNSPGKYLEFFPAGNKNPEYVGAFSYEVLTVE 295
QY 403 NTOAERLRGLVQOR-----EGVIISASQEOIRELTRDSESRMHIRRGSGSSRP 454
DB 296 NTQREKLEELIEQQRQKQOQOQGMFRKAKPEQIRAIQOATSPR-----HNGGE--RLA 349
QY 455 YILFNKRLPKSKYGOAYEVKPEDYROLQDMVSVFIANIITGOSMMGPFNFRSTRVVV 514
DB 350 INLLSQPYVSNQNGFEFFACPEDFSQFQNMVAVASAFRLNGALFVPHYNSKATFFVYV 409
QY 515 ASGEADVACPHLSGRHGRGK--RHEEEDV-----HYQYKARLSKREAIIVPV 566
DB 410 TQGYAGQACPHLSROSQSGOSGRQDRREDEESEEFTFGEFOOVKAPLSGDVFAVA 469
QY 567 GHPVYFVSSGNENLLFAGFINAONNHNENFLAGR 600
DB 470 GHAVTFPASKDQPLNAVAFGLNAONNQRITFLAGR 503
```

RESULT	7			
ID	003865	PRELIMINARY;	PRT;	582 AA.
AC	003865;			
DT	01-NOV-1996	(TREMBLrel. 01, Created)		
DT	01-NOV-1996	(TREMBLrel. 01, last sequence update)		
DT	01-OCT-2000	(TREMBLrel. 15, last annotation update)		
DE	VICILIN-LIKE EMBRYO STORAGE PROTEIN.			
GN	GUB1-L.			
OS	Zea mays (Maize).			
CC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
CC	Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.			
OX	NCBI_Taxid=4577;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=M6A;			
RX	MEDLINE=92090707; PubMed=1752424;			
RA	Belanger F.C., Kriz A.L.;			
RT	"Molecular basis for allelic polymorphism of the maize globulin-1 gene.";			
RL	Genetics 129:863-872(1991).			
DR	EMBL; X59083; CAA1809.1; -.			
DR	HSP; P50477; ICAU.			
DR	MENDEL; 11234; Zeama:1188;11234.			
DR	INTERPRO; IPR000901; -.			
DR	INTERPRO; IPR001113; -.			
DR	PFAM; PF00546; Seedstore_7s; 1.			
DR	PROSITE; PS00867; CPAS_2; UNKNOWN_1.			
DR	PRODOM; PD081059; -; 1.			
DR	Seed storage protein.			
Q	SEQUENCE 582 AA; 66162 MW; 680D85FECD7CE885 CRC64;			

Query Match	25.3%	Score	895;	DB	10;	Length	582;
Best Local Similarity	39.6%;	Pred. No.	5.8e-57;				
Matches 201; Conservative	83;	Mismatches	183;	Indels	40;	Gaps	11

QY	182	EDNRKRPDORREYEDCSRRCQDQSRPOYOOCQRRCRE-----QORHNGSGDILINORG	259
Db	25	EDDHHHHHGGKSSQCYRCEEDRHMWRPRLCEDQKREERKROERSRHAD-----RS	79
QY	236	GGG---RYEBGEERKQSD-NPYFDEBSLSTRPTEEGHISVLENFYGRSKLLRALKNR	230
Db	80	GGGSEDEDEREOKCKQDRRPVYEDRRSPRVAVRSFEGSGSLRYLRPFDEVSRLTGLGIDYR	139
QY	291	LVLEANDPNAFVLPETHLDALAILLVYSGRALKMHINDNEESVNLBEGGDVIRIPAGTTFY	350
Db	140	VAVLEANDRSVNVSHTDHACICVABEGGVTTTENGERRSTYIKQGHFAVPAGAVTY	199
QY	351	LINDNNEERLHIAKELQTIISTPGYKFEFPAGGONPERPYLSTFESKELLEALMTQOERLR	410
Db	200	LANTGKRKKLVITKLTITISVPGFQFFGFGCGNRPESFLSTFSKSIQORAYATSSDRLE	259
QY	411	GVLGO--QREGVIIISASQEOIRLETRODSE---SRHHNIRGGSESGRPNLNFKRPLYS	465
Db	260	RLFRHGOQDGILVRAATEEOTRELRRAHSGHGHPMLRPFPGE-SRGPYSILDRPESIA	318
QY	466	NKYGOAEYKRPEDYROTQMDIVSYFVFINITOGSMGCGFMTNSTKVVAVSGEADVPMAC	525
Db	319	NQHQLTEADARSHDLAEHDVSYSPFNITAGSMSAFLPVTREKTIAYVNGKGYAEIYC	378
QY	526	PHLSGRHG---GRGGRKHDEEDVDV-----YEOVKARLSKREAIIVVPVCHPVV	571
Db	379	PHROSQGESEEREDMKGRSEEEDEEESSEBOEAGQCYHTIRARLSGTFVVPVAGHPV	438
QY	572	FVSSGNELLLFARGINONNHHEFLVGRENNYLOOEPALEALFAPKAEVEELFENSO	631
Db	439	AVASRDSNLOIVCEVYADNREKLYFLAGAD-NYLQDKDRAKALSFASKAEVDEVLGSR	497
QY	632	DESTFFPQPMQ--HQOQSSBSTKQOQ	655
Db	498	KEKGFGLPGRPESSGHEERDEEEREE	524

RESULT	8		
ID	003678	PRELIMINARY;	PRT; 637 AA.
AC	003678;		
DT	01-NOV-1996 (TrEMBLrel. 01,		Created)
DT	01-NOV-1996 (TrEMBLrel. 01,		Last sequence update)
DT	01-OCT-2000 (TrEMBLrel. 15,		Last annotation update)
DE	FETAL STORAGE PROTEIN.		
GN	BGLI OR GBLI.		
OS	Hordeum vulgare (Barley), and Triticum aestivum (Wheat).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
CC	Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.		
OX	NCBI_TaxID=4513, 4565;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=93287988; PubMed=8510647;		
RT	Heck G.R., Chamberlain A.C., Ho T.H.D.;		
RT	"barley embryo globulin 1 gene, Bgl1: characterization of cDNA,		
RT	chromosome mapping and regulation of expression."		
RL	Mol. Gen. Genet. 239:209-218(1993).		
DR	EMBL: M64372; AAA32936.1; -;		
DR	EMBL: M81719; AAA34265.1; -;		
DR	HSSP: P02853; 2PHL.		
DR	MENDEL: 8553; Horvu_1188; 8553.		
DR	INTERPRO: IPR000901; -;		
DR	INTERPRO: IPR001113; -;		
DR	PFAM: PF00546; Seedscore_7s; 1.		
DR	PROSITE: PS00867; CPSASE_2; UNKNOWN_1.		
DR	PRODOM: PD081059; -, 1.		
KQ	Seed storage protein.		
SQ	SEQUENCE 637 AA; 72252 MW; F323MFFF99947C3C CRC64;		

Query Match	24.8%;	Score 875;	DB 10;	Length 637;
Best Local Similarity	35.6%;	Pred. No. 1.9e-55;		
Matches 227;	Conservative 99;	Mismatches 206;	Indels 106;	Gaps 22

[illegible]

```

Db 460 PHLGSESEEREHGK--GARREEDDDQOORRRGSESESEEEEOQRVETVARVSRGS 517
QY 561 AIVYVGVGHVYVVS--GNEMLLFVFGINAOHNHNFVLAGRRNVLOQIEPAMELATA 618
Db 518 AFVYVPGHVVIVSSQGSNLOQVCFELIAENNERVWLAGR--NVVIGKLGSPAOLFTFG 576
QY 619 APRKEVEELFNODEST--FFPGPROHOOGSSRSTKOOO 655
Db 577 RPARVQVYFRAQDDDEGFVAGP----EQOSREOEOEQ 610

RESULT 9
O9M3X6 ID O9M3X6 PRELIMINARY: PRT: 613 AA.
AC O9M3X6:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CONVICTILIN PRECURSOR.
GN CVC.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Pisum.
RN NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BIRTE; TISSUE=SEED COTYLEDON;
RA Casey R.;
RT "The sequence of a pea convictilin cDNA."
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ276875; CAB82855.1; -.
KW Signal.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 613 CONVICTILIN.
SQ SEQUENCE 613 AA; 72063 MW; 49AED99F6135DD19 CRC64;

Query Match 24.1%; Score 851; DB 10; Length 613;
Best Local Similarity 30.9%; Pred. No. 9,7e-54;
Matches 212; Conservative 135; Mismatches 236; Indels 104; Gaps 14

QY 1 MAINSNSLCILFLSLFTSLFTTYSVLASEPFROGEYEECKRQCMQLETSGCMRRCSQCD 60
Db 1 MATTIKSRPPLLGLITFLASVVCVTYANYD-----EGSEPRVPAQRRGHEE 51
QY 61 KRFEEIDIM-SKYNDODPOTDCQOCORRCROESGPRQOQYCOR---CKEICEEEE 115
Db 52 KEKRGHGMRSYKEEDE---EGQNRGNGQEGBEKERRKGEMRPSYKEDDEEEKK 107
QY 116 YNRQDPOQOYEQOCORRHETPRHMOTCOQRCERYEKKRKOQKYEEQOREDEEK 175
Db 108 YRYOREKDEEEKOYQYQOREKKEQVOPGRERWEREDEQVDVEWGSORREDPEER 167
QY 176 YEENKEDDNRRDPQ-ORVEYEDRCRRCOQEPERQOYQOCRRREQORRGROGDLINPQR 234
Db 168 ARLNRERRTKDRRHQXEGEEESSSESQER----- 200
QY 235 GSGRGYEEGEEKOSNDPYVFDERSLSTRFTEEGHISYLENFYGRSKLLRALKNRYLVLL 294
Db 201 -----NPLFKSNKRLTLTFENNNGHIRLQORDKRSDLEFNLONRVLEY 245
QY 295 EANPNAYVLPFHLDADAILLVYTGAGALKMIHRDNRESYNLECGDIVIRIPAGTFEYLNR 354
Db 246 RAKPTTIFLPQHIDADLLLVYLSKAILTVLSPNDRNSYNERGDTIKLPAGTTSYLVNQ 305
QY 355 DNNEGLHIAKFLQITSTGQYKPEFPAGQONPEYPLSTFSKILLAAALTOQERLGVIG 414
Db 306 DDEEDRLVLDVIVPVGNGGKEAFELA--KKNQYLRGFSKNILASVYTRJETIEKVL 363
QY 415 QQRE-----GVYISAOQIREFLTDDSDSRWMIHTRGGESSRGYVFNLFNPKR 462
Db 364 EEQEKDRKRRQOGEETDAIVKYSRQIEELKKLAKSSS---KKSIPSEFEIINLRSHKP 419

```

OY	463	LVSNNVGGAYEKKPE-DYRQLODMOVSVLIANTIGSGMSPFENFTSTVVVAASCEADV	521
Db	420	EKSNNFGKIFETTPTEKKTYQDDDLFLFVSCVEINEGALMLPHYNSAIYVLVLLNEKGKL	479
OY	522	EMACPLSLGRHGRCGRGKRHEEBEE-----DVHYEQVARLSKREAIIVPVGHVFEVS	574
Db	480	ELLT-----GLTKNEQOEREDRKERRNNEVQRYEARLAPCPDVIILPAGHPAATA	526
OY	575	SENENULLFAFCINNQNHNENFLAGERNVLDOIEQAOMELFAAPRKVELELFNSODDS	634
Db	527	SSNLML-L-GGEINAMENNRNFLSSDDVDISOIENPAVELTFPPGSVOEIINRLINKOKS	584
OY	635	IFFPGPRHQHOQSRSRTKOQOLPVASIL	661
Db	585	HFNANPECKEGESOG--KRPSSLSTIL	609
RESULT	10		
ID	022120	PRELIMINARY:	PRT; 544 AA.
AC	022120:		
DT	01-JAN-1998 (TREMBLrel_05 Created)		
DT	01-JAN-1998 (TREMBLrel_05; Last sequence update)		
DT	01-OCT-2000 (TREMBLrel_15; Last annotation update)		
DE	ALPHA SUBUNIT OF BETA CONGLYCININ.		
OE	Glycine max (Soybean).		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;		
OC	Fabales; Fabaceae; Papilionoideae; Glycine.		
OX	NCBL_TaxID=3847;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. WASESUZUNARI;		
RA	Maryama N., Katsube T., Mada Y., Utsumi S.;		
RL	Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AB008678; BAA23360.1; -.		
DR	HSSP; P50477; ICNU.		
DR	MENDEL; 25074; GLYMA:1188;25074.		
DR	INTERPRO: IPRO01113; -.		
DR	PFAM: PF00546; Seedstore_7s; 1.		
DR	PRODOR: PD081059; -.1.		
SQ	SEQUENCE 544 AA: 63296 MW; 048B243929A9C9AB CRC64;		
<hr/>			
Query Match 23.6%; Score 833; DB 10; Length 544;			
Best Local Similarity 32.3%; Pred. No.1.7e-52;			
Matches 186; Conservative 135; Mismatches 195; Indels 60; Gaps 12.			
OY	107	KICEEEE---EYNQRDPQQQYEDQCERCORHNETEPRIHMOTCOQRCERYEKERKQOK	163
Db	4	KEECEGETPRPRRHQPHREPOQGKEEDEDEDQPRIPRRPROPQREEHDEQEDE	63
OY	164	RYEEOQRDEDEYEERMKEEDMKRPDQOREYDCRRRCQEQOPRPOQYQCARCRREQORH	223
Db	64	WPKREKKGKESGEDEDEDEDERQPFP--RPHQKERRQOEDED--EEDQIRS	118
OY	224	GRGGDLINDORGSGSYEEGEKQ-----SDNPYFDERSLSTRTEREGHTSVLENFYG	278
Db	119	-----BESEDESLRRHKKNKPFLFGSKRFETLKKNQYGRIRVLQREFNQ	161
OY	279	RSKLLRALKNRVLLLEANPNAFVLPYLHDADAILLVTGRCALKMIHRDNRESYNLBCG	338
Db	162	RSPQLONLRDRIETLEFNKSKPTLLPHNHADADYLVINGTAILLVNDDRDYSRLQSG	221
OY	339	DVIRIPACTTFFLLMRDNNERLHIKFLQITISPCQYKEFPFAGQGNDEPYLTSTSKEL	398
Db	222	DALRVPSGTTYYVNDPNENNLRLLTLAI PVNKPGFRFSFLLSTEAOSYLDQSFNLL	281
OY	399	EALULTQAEIRLGVL-----GOQR--EGVITSASOEQITRETTRDSEGRPMHIRG	447
Db	282	EASYTRKEELINKVLFSEDEGGQGGEQRLQDSYVILEIKBOIRALSKRAKSSSKRTI---	338

Qy	448	GSSSGPNLLEFKRPLYSNKKGAQAEVKEPEYRQLOQMDVSVFLANTIGQSMMPFENR	507
Db	339	--SEDKPFLKSRDPLYSNKKLGKFFETLPKRNPLRDLIDFLSTVDNMGALLPLPHNSK	397
Qy	508	STRKVVVAVSGEADVMAACPHLSGRHGG--RRGGRKHEEEDVHYEQVKARLSREALVVP	565
Db	398	ALVILVINEGDANIELV-----GLKEQQQEQQQEQEQEGLVRYKTRALSSBDIDVIP	448
Qy	566	VGHPEVFPVSSCGNELLLEFAGIMANONHNENLAGRERNVLOQITEPAMELAFAPAKREVE	625
Db	449	AGYPVVV--NMTSNLNFPAIIMENNORNLVLAGSDNVLSQIPSOVELAFPGSAQAVE	506
Qy	626	ELFNSODESIFPPGPRHQQOOSRSTKQOQPLVSLT	661
Db	507	KLLKNQRESYFDAPKKKEGKNGKR--GFLSSTL	540
RESULT	11		
	041674		
ID	041674	PRELIMINARY;	PRT; 545 AA.
AC	041674:		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)	
DE	CONVICITILIN PRECURSOR.		
OS	Vicia narbonensis.		
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;		
OC	Manuoliophyta; eudicotyledons: core eudicots; Rosidae; eustosids I;		
OC	Fabales; Faboaceae; Papilionoideae; Vicia.		
OX	NCBI_TaxID=3912;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=CYTLEDON;		
RA	Koch G., Koenig S., Becker C., Horstmann C., Schlesier B.;		
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: Z71986; CA96513.1; -.		
DR	HSSP: P50477; ICAU.		
DR	MENDEL: I2432; Vicia:1188;12432.		
DR	INTERPRO: IPR001113; -.		
DR	PFAM: PF00546; Seedstore_7s; 1.		
DR	PRODOM: PD081059; -; 1.		
KW	Signal: Seed storage protein.		
FT	SIGNAL	1 28	POTENTIAL.
FT	CHAIN	29 545	75 GLOBULIN, CONVICITILIN.
SO	SEQUENCE	545 AA; 62810 MW; 459A876F92F5A87E CRC64;	

	Query Match	23.2%	Score 818;	DB 10;	Length 545;
	Best Local Similarity	34.3%;	Pred.No. 2.1e-51;		
	Matches 196;	Conservative 101;	Mismatches 204;	Indels 70;	Gaps 14.
Oy	109 ICEEEBVRNRDPOOQVOCOCRCORHTEPRHMOTCOOCERRYEKERRKQRYEE-	167			
Db	23 LCVTANYANDEGTETPRVPGGRERGROGEKEKRH-----GEMRSHIEKAQPGREKM	75			
Oy	168 QOREBEKYEEKMKEDNKRDPQOEYEDCRRCEQOEPRQOYOCCRRCEQOROHGGG	227			
Db	76 ETSEEBEERVDENWRKSQRHEDEERARE--RRAEERERRRQWE-----	117			
Oy	228 DLINPORGSGRGYEEGEEKOSD-NPYEFDEKSLSTRFTEGHISVLNFYGRSKLRA	285			
Db	118 -----GEEKSSKSSQGERNPFLFKSMKLTLPFNENGHIRRLRFPKRSLLFFEN	167			
Oy	286 LKNYLVLYLEANNPAFVLFTHLDADAILVYGCGALMKTIRHDNRESYLVECGDVIRIPA	345			
Db	168 LONYLAVERYAPRHILFLPHOIDADILLIVLSGRAILTVLSPDRNSYMLLEGDTIKLPA	227			
Oy	346 GTTFPLINDNNERTHIKAFLOTISTPGQYKEFPFAGGONPEPVYSTFEKELLEAALNQ	405			
Db	228 GTTSTLLNODDEDLRVYDSLTSYNRPKVESFELGSKN-QYLRGFKNILLEASLNKK	285			
Oy	406 AERLGVL-----GOORE-----GVIIASQEQIRELTRDSESRRWHIRIGESS	451			

```

Db      266 YETIEKVLLEERPOOSIGCKRRKSQOETNALVKYSREOVEELKRLAKSS-----KKGVSE 341
OY      452 RGPYLLFNKRLPYLSNKYQOAYEVKPE-DYROLQMDVSVFIANTTQSGMMGPFENRTSTK 510
      1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 :
Db      342 FEPFLRLRQNKRSYKFNKLELPEKPKYDLOLODLFISSVETINEGGLMIPHNSRAIV 401
      1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 :
OY      511 VVVVVSAGVDYEMACPHLSGHRHGRGGKRRHEEDVHYEDVKRLSKREALIYVPCHPV 570
      1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 :
Db      402 ILLVNEGGLNE-----LVGLKNEQOQREEREDHQV--QRYEARLSPGDVILLIPACHPV 454
      1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 :
OY      571 VVVSSGNLLELLFAFGIAQNNHNNHNFVLAGRRNRVLOOIEPAMELAFAPARKVEYEELFNS 630
      1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 :
Db      455 AVSASSNLTLL--GFGIAVENNRNRFLETGSDVDNISQIENPVKELTFPGSAQEVNRLKN 512
      1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 :
OY      631 ODESLTFPPGRQHQOOSRSKTKQOQPLVSTL 661
      1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 : 1 :
Db      513 QEHSHF--ANAEPEQKGEESQKRSKSPSTSL 541

```

RESULT	12			
Q9SP11				
ID	Q9SP11	PRELIMINARY;	PRT;	489 AA.
AC	Q9SP11;			
DT	01-MAY-2000 (TREMBLrel_13, Created)			
DT	01-MAY-2000 (TREMBLrel_13, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel_15, Last annotation update)			
DE	SUCROSE BINDING PROTEIN HOMOLOG S-64.			
GN	SBP.			
OS	Glycine max (Soybean).			
OC	Eurycotylae: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:			
OC	Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids I;			
OC	Fabales; Fabaceae; Papilionoideae; Glycine.			
OX	NCBI_TaxID=3847;			
	[1]			
RN	SEQUENCE FROM N.A.			
RP	Pedra J.H.F., Delu-Filho N., Pirovani C.P., Contim L.S., Dewey R.E.,			
RA	Otoni W.C., Fontes E.P.B.;			
RT	"Antisense and sense expression of a sucrose binding protein homologue			
RT	gene from soybean in transgenic tobacco affects plant growth and			
RT	carbohydrate partitioning in leaves."			
RL	Plant Sci. 0:0-0(1999).			
DR	EMBL: AF191299; AAF05723.1;			
DR	HSSP: P02853; 2PHL.			
DR	INTERPRO: IPR001113;			
DR	PFAM: PF00546; Seedstore_7s; 1.			
SO	SEQUENCE 489 AA; 55834 MW; 9BBC0D45EDECCD2 CRC64;			

Query Match	22.6%	Score 798.5	DB 10	Length 489
Best Local Similarity	35.4%	Pred. No.4.6e-50		
Matches 173	Conservative 99	MissMatches 172	Indels 45	Gaps 11
OY 179	RMKEEDNRRP-----	-OQREYED-----	CHRRQCGQRRPROOYOQRCRCREQROH	223
Db 27	KKEFEVEDEPRLVYCKKIDOCQOORITTEDSKTKCLQOCCSMQOEKYOVEEETRKEEKH			86
OY 224	GGGGLINDPORGSGRYEGEGEKQSDNPYPDE-RSLSTRFTEGHSIVLNFENYGRSKL			282
Db 87	-----QDRHEEQDQONPVVFEDEDDFSTRVTEGGSTRVLKLFTEKSKL			130
OY 283	LRLAKNRYRLVLEANPNNAVLFTHLDDADILLVYGGGALKMIHRDNRESYLECGDYIR			342
Db 131	LOGNENFRATLEAAHRAHFVSPRHDESEVLLENIGRAVLGLARESETEKITLEPGDMTH			190
OY 343	IPAGTFFVLINRDNERRLIHAKFLOTISTPGQYKKEFFRPAQGNPPEPLYLSTPSKELLEAL			402
Db 191	IPAGPLVIVNDEDEKILLALMLHPVSPCKFEFFFGGGGDPDSVYSAFPMNVLOAL			250
OY 403	NMQAERLNGVLCQOREGYTISASQOIRBELTDDSESSRWMTIHRGEGSSRGYNLFENR			462
Db 251	OTPKRLERLFNQDNEGSLFKISRVRALA-PTKXSSWMP--GGE-SKAOFNLFISARP			306
OY 463	LYSNRYGQAYEVRKPDYRQ-LQDDMVSYFIANTGSGMNGPFFNTRSTKRVVVAAGADV			521

Db 307 TFSNGGRLETEVGPDDEKSWLQRLNMLTFTNTQSRMSSTIHNSHATKALATMDRGHL 366
522 EMACPHLSGHRGGRGKRHEEEDVHYEQVKARLSKREAIYVPGHPVYVSSGNENLL 581
Db 367 QISCPHMSRSBSK-----HDKSSPSYHRISADLPGWVFPVPPHPYTTASKNENLL 420
582 LFAFGINAONNHENFLAGRENNVLOQIEPQAMELAFAPRKEVELLFSNDSIESFFPGPR 641
Db 421 IICEFVNVADNKKFTFAGKD-NIVSSLDNVAKELAFNPSEMVNGVF-ERKESLFFPEFEL 478
Qy 642 QHQOQSSRS 650
Db 479 PSEERGRA 487

RESULT 13
ID 049927 PRELIMINARY: PRT: 483 AA.
AC 049927;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE P54 PROTEIN.
GN P54.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RL Castillo J., Marquez J.A., Franco L., Ballesstar E., Rodrigo M.I.;
RA Submitted (Feb-1997) to the EMBL/Genbank/DBJ databases.
EMBL: Y11207; CAA72090.1; -
DR HSSP: P50477; ICAU.
DR MENDEL: 28319; Pissa: 3005; 28319.
DR INTERPRO: IPR001005; -
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore-7s; 1.
DR PROSITE: PS00041; HTL_ARAC_FAMILY_1; UNKNOWN_1.
DR PRODOM: PD081059; -; 1.
SO SEQUENCE 483 AA; 54662 MW; 8127BDAAD178F3D CRC64;

Query Match 21.8%; Score 770.5; DB 10; Length 483;
Best Local Similarity 31.6%; Pred. No. 4.9e-48;
Matches 163; Conservative 113; Mismatches 169; Indels 71; Gaps 12;

Qy 135 RHETPRHMQTCQQRCEKRYEKERKQKRYEQOQREDEDEKTEENKEDNKRDPQOREY 194
Db 28 RREKDP-ELTTCKDQCD-----MQRQYDEDEK----- 53
Qy 195 EDCRRRCQEQEPRQOYQOCRCQEQORHGRGDDLNPORGSGRYEEGEEKOSNDPYF 254
Db 54 RICEKCD-----DIKKKQERQKHKEH-----EEEEQEQDEDEPPYF 93
Qy 255 DERSLSTRTEEGHISVLENFYGSKLRLALKNYLLVLEANPAFVPTHLADAILL 314
Db 94 EDNDEFTKIDTIDGKYLINKEKNEKSKLLKNIENTENGLAVLEIKANAFSLPHYDSAILF 153
Qy 315 VNGRGALMKIHRDNRESYNLECGDVIPIAGTFFYLINRDNNEHLIAKF--LQITST 371
Db 154 NIKGRGIIQLVADERTERENLEEGDIMRVAGTPTVLVRODEKELYIAFIHPPSSGA 213
Qy 372 PQOYKEFPAGQONPEPYLSTFSKELLEALNTQAERLGVLGQOREGYITASQOIRE 431
Db 214 PVNLEFEFFSARKEPSVLTSSKVLQALAKSSKGELETVLDEOKGRIFIKEDVAG 273
Qy 432 LTRDSESRHMIIRGESSRGPYNLFNKRPLYSNKYGAYEVKPPDYQ-LQMDVSVF 490
Db 274 LAKPKS---LWPF---GGFFKSPFNIFSNPAFKNKFGSLFEVGSQEKSGLEGLMLLT 327

Qy 491 IANITGSMKGFENFTRSKRVVVASGEADVEMACPHLSGRHGRGKRHEEEDVHYE 550
Db 328 LANITGSMSTIHTNNAKNALIVIDGELEMACPHMSSSSNSR-----QKSSISYH 382
Qy 551 QYKARLSKREAIYVPGHVVYVSSGNENLLFAFIGNONNHENFLAGRENNVLOQIEP 610
Db 383 MNAKLPGWVFPVPAHGHPFNATASKKNLIVCFVNAQNRKNLALAK-KNIVSALDK 441
Qy 611 QAMELAFAPRKEVELLFSNDSIESFFPGPRHQOQ 646
Db 442 AAKEVAFDIAAEKVDYEFERKE-FFFPYDNEERKE 476

RESULT 14
ID 040873 PRELIMINARY: PRT: 448 AA.
AC 040873;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE VICILIN-LIKE STORAGE PROTEIN.
OS Picea glauca (White spruce).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Picea.
OX NCBI_TaxID=3330;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PG118; TISSUE=SOMATIC EMBRYO;
RX MEDLINE=93004485; PubMed=1391775;
RA Newton C.H., Flinn B.S., Sutton B.C.;
RT "Vicilin-like seed storage proteins in the gymnosperm Interior spruce
RL Plant Mol. Biol. 20:315-322(1992).
EMBL: X63191; CAA44873.1; -
DR HSSP: P02853; 2PHL.
DR MENDEL: 12115; PICGL1188; 12115.
DR INTERPRO: IPR001113; -
DR PFAM: PF00546; Seedstore-7s; 1.
DR PRODOM: PD081059; -; 1.
KW Seed storage protein.
SO SEQUENCE 448 AA; 50200 MW; 467AE5BA12E6D157 CRC64;

Query Match 21.7%; Score 765; DB 10; Length 448;
Best Local Similarity 37.8%; Pred. No. 1.1e-47;
Matches 162; Conservative 89; Mismatches 154; Indels 24; Gaps 9;

Qy 232 PQRGSGRYEEGEEKOSNDPYFDEKSLSTRRTTEGHSVLENFYGSKLRLALKNYRL 291
Db 34 PEYLGGRGR-EEEREENPVYFHSDFTRASSSEGETRALPNFGEVSELLEGIRKFFV 92
Qy 292 VLEANPAFVPTHLADAILLVYTGREGALMKIHRDNRESYNLECGDVIPIAGTFFYL 351
Db 93 TCLEMPKNTVMLPHYTDAILVYTRGRYIAYVHONELYKRLLEGVDFGVPSCHTFL 152
Qy 352 INKDNNEHLIAKFLQITST-PGOYKEFPAGQONPEPYLSTFSKELLEALNTQAERL 410
Db 153 VNNDHNTLRISLVRPVSTVGEQPFYVAGGRNQTYASAFSDVLEAANTVNOQLE 212
Qy 411 GVLGQOREGYITASQOIRELTRDSESRHMIIRG--GESSR---GPNLNEKRL 463
Db 213 RIFGGRKSGVITHANDEQIREMR-----KRGSAGSMSAPRHPKPNLNRQKD 262
Qy 464 YSNKGOAEEVAPEDYRQDMDVSYFINITNGOSMGSPFTNRSTKYVYVASSGEADVEM 523
Db 263 FENENGRFTIAGPKNPTFDALDVSYGLADLNPGSTWTAJSLNKSSTISGITVNGGRLEM 322
Qy 524 ACPHLSGRHRGGRGKRHEEEDVHYEQVKARLSKREAIYVPGHPVYVSSGNENLLF 583
Db 323 ACPHL-GQHGW---SSPREGDDDIYQVMAKLRTGYSYIYVAGHPITFIATSNRLQIL 379
Qy 584 AFGINAONNHENFLAGRENNVLOQIEPQAMELAFAPR-KEVELELFSNDSIESFFPGPR 642

Db	380	MFDLNTRNENQFLAGK - NNVLNLTLEKEIRKLSNVRGEEIEFLVLOAQKDQVILRGFQR	438
QY	643	HQQSSRST	651
		: : : :	
Db	439	RSRDEARSS	447

Job time: 1557 sec

RESULT	15			
022121				
ID	022121	PRELIMINARY;	PTT;	417 AA.
AC	022121;			
DT	01-JAN-1998	(TEMBLrel, 05, Created)		
DT	01-JAN-1998	(TEMBLrel, 05, last sequence update)		
DT	01-OCT-2000	(TEMBLrel, 15, last annotation update)		
DE	BEA SUBUNIT OF BETA CONGELYCININ.			
OS	Glycine max (soybean).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;			
OC	Fabales; Fabaceae; Papilionoideae; Glycine.			
OX	NCBI_TaxID=3847;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CV. WASESUZUNARI;			
RA	Moriyama N., Katsube T., Wada Y., De La Rosa A., Usumi S.;			
RL	Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AB008679; BAA23361.1; -.			
DR	IIISP; P02853; 2PIL.			
DR	MENDEL; 25075; Glyma;1188;25075.			
DR	INTERPRO; IPR001113; -.			
DR	PFAM; PF00546; Seedstore_7s; 1.			
DR	PRODOM; PD081059; -; 1.			
SO	SEQUENCE	417 AA;	48106 MW;	DDEE5085B779556E CRC64;

Query Match	21.4%	Score 757;	DB 10;	Length 417;
Best Local Similarity	37.6%	Pred. NO.	3.8e-47;	
Matches 160;	Conservative	89;	Mismatches 150;	Indels 26;
				Gaps 7;

QY	246	KOSNPYVF--DERSLSTRFRTEEGHISVLEENFYGRSKLLPALKNNRYLVLEAVNPVFLP	304
Db	6	EDENNPFYSSNSFOYLFENONORITLLOFRKRSQLEMLADRYLVQYQSRPNITLLP	65
QY	305	THLDADAILLVATGGAGLKMTHRDNRESYNLECGDVIRIPAGTTFYVLINRDNNERLHIAK	364
Db	66	HHADADFLLEYSRALLTVLNNDDBRSYMLHFGDAORIPAGTTFYVLNPHDQNLKTIK	125
QY	365	FLQIISIPGQYKKEFPAGGQNPPEYLTSEKTELEALNTQARLRGLV-----GQORE	418
Db	126	LAIPVNPGRIRDDFELISTQAGSYLOGFSHNILLETSHSEFEELIRKTVLFGEEEDOROE	185
QY	419	GVIIASAQEOILRELTDRDSESRMHIRGSESSRGPNLENRKPLYSNKYQAYEVPKD	478
Db	186	GVILELSKEQIRQLSRAKKSSSKRTI-----SSEDEPNLRSKRPITSNNGKFFELTPK	241
QY	479	YROLQMDVSVFLANITQSGSMGCPFFNTSTKVVVVAASGEADVEMACPILSGHHG--RR	536
Db	242	NPOLRDRDIFLSSVDINEGALLPHFNSKAIYLVILINEGANIELV-----GIKED	292
QY	537	GGRHEEEDVHYEYQVARSLSREALVYVRGHHVVPVSSSNEULLPARGINNQNHNEN	596
Db	293	QOKKOQEEBPELVORYRAELSEDDVFLIAPAY--FLVNTYSMLNPLAGINENNOQRNF	350
QY	597	LAGEERNVLOQIEPOMELFAAPRKVEVELFNQSDSEIFFPGRQHOQSSSSTKQOOP	656
Db	351	LAGKMDVVRQIEQYQELAFPGSADQVERLLKKQKRESYVYDQPOQKBEQSGGRK--GP	408
QY	657	LVSTL 661	
Db	409	FPSTL 413	

